EXHIBIT A

Supplemental Report – Neurontin
Sheila Weiss Smith, PhD FISPE

My name is Sheila Weiss Smith, and I am a Professor in the Departments of Pharmaceutical Health Services Research, School of Pharmacy and Epidemiology & Preventive Medicine, School of Medicine, University of Maryland Baltimore. Currently, I am a visiting scientist at the National Institutes for Health (NIH). I also hold appointments at the Veteran's Administration (Research Associate) and the Department of Epidemiology, Johns Hopkins University's Bloomberg School of Public Health (Visiting Professor). I am a fellow of the International Society of Pharmacoepidemiology. Over the past decade, I have served as a voting member on a number of advisory committees, when epidemiologic expertise was needed. My expertise and professional accomplishments are more fully documented in my original expert report dated: December 20, 1997. Since my last report, I was promoted to Professor and began a one-year sabbatical at NIH.

At the request of counsel, I have prepared this supplemental report, which sets forth my opinions concerning FDA's Statistical Report and Evaluation on Antiepileptic Drugs and Suicidality (May 23, 2008), and the subsequent meeting of members of the Peripheral and Central Nervous Systems and Psychiatric Drugs Advisory Committees (July 10, 2008) on the topic of antiepileptic drugs and suicidality, as well as certain opinions expressed in reports and Declarations prepared by other experts that have been filed since the date of my original report. Based upon my review of the available data it is my opinion that Pfizer has correctly determined that there was no signal of suicidality associated with Neurontin. The FDA's post-hoc analysis of prospective randomized placebo-controlled clinical trial data from a select group of 11 antiepileptic drugs (which included Neurontin) has several methodological flaws. FDA's analysis does not provide evidence that Neurontin is associated with suicidality and it certainly does not provide evidence upon which causality can be established for this endpoint. Based on my review of the data, it is my opinion that the Neurontin package insert appropriately reflects the data, which do not suggest the need for a warning or precaution regarding suicidality.

I hereby adopt and incorporate my earlier work in my original expert report dated December 20, 1997. In this Supplemental Report, I am responding to opinions and testimony of plaintiffs' experts and FDA's statistical review and evaluation of suicidality with antiepileptic drugs that were expressed subsequent to my first report. I also offer opinions regarding Pfizer's conduct in the development, testing and labeling of Neurontin, which were not part of my general causation opinions in my first report. The opinions expressed in this Supplemental Report are held to a reasonable degree of scientific certainty, and are based on my education, training, experience, and my review of the relevant scientific and medical literature. I have reviewed the supplemental expert report of Dr. Robert Gibbons, and I adopt and incorporate by reference the contents of that report. Additional materials considered are attached as Exhibit 1 and copy of my current CV is attached as Exhibit 2. I reserve the right to review and rely upon subsequent literature and reports filed by other experts.

Opinion 1. FDA conducted a meta-analysis of clinical trial data across 11 antiepileptic drugs. In my review of FDA's report, the transcript of the FDA Advisory Committee meeting, and other related materials, I found that there was inadequate bases to reliably assert that Neurontin was associated with suicidality. Pfizer was correct in concluding that based upon the available data, there was no signal for suicidality with Neurontin.

The FDA asked the sponsors of certain marketed antiepileptic drugs to provide data from randomized controlled clinical trials that met specific inclusion criteria. (Katz R. 2008) It is not clear from the report how many companies were queried and how many did not respond. The analysis was based on eligible clinical trials for a total of 11 antiepileptic drugs; carbamazepine, divalproex sodium, felbamate, gabapentin, lamotrigine, levetiracetam, oxcarbazepine, pregabalin, tiagabine, topiramate, and zonisamide. (Levenson M et al. 2008) Of these drugs, three had labeled warnings about an elevated risk of suicide or suicide attempts based on elevations in suicide attempt rates seen during clinical trials. These are levetiracetam (4 patients on treatment (0.5%) versus 0 on placebo attempted suicide), topiramate (13 patients 0.3% on treatment and 0 on placebo attempted suicide), and zonisamide (1.1% on treatment and 0.4% on placebo attempted suicide). (ORTHO-McNeil 2006, UCB, Inc 2006, Eisai, 2007) The lamotrigine label has a general precaution about the clinical worsening and suicide risk associated with bipolar disorder. (GlaxoSmithKline 2006)

I agree with Dr Gibbons' explanation of the utility of meta-analyses such as the meta-analysis of clinical trial data of 11 anti-epileptic drugs prepared by the FDA. A meta-analysis is essentially an analysis of analyses: it is a statistical methodology that analyzes a large collection of results of individual studies for the purpose of integrating the findings. Meta-analyses are "useful methods for research synthesis, for taking a lot of information and pulling it into one general statement that may say something about the overall effects." (Gibbons Daubert Hearing, Day 2, June 20, 2008, at 298:24 – 299:2) It is important to understand that a meta-analysis is not a prospective randomized clinical trial. Rather, it is a "tool for research synthesis" applied retrospectively to prospective randomized clinical trials. (Gibbons Daubert Hearing, Day 2, June 20, 2008, at 300:4, Berlin and Kim 2005, Chapter 44). Meta-analyses "are not useful for deriving causal inferences. (Gibbons Daubert Hearing, Day 2, June 20, 2008, at 299:2 – 3). This is especially true when there is considerable heterogeneity among the trials in a meta-analysis (Thompson SG. 1994), as was clearly evident in the FDA study. Some examples of heterogeneity include, but were not limited to, the diverse group of agents studied, a large and diverse accumulation of indications, and a wide age range of patients (pediatrics to elderly). (Levenson M et al. 2008)

A key to performing a proper meta-analysis is accounting for the heterogeneity of results of the various individual studies. According to plaintiffs' expert, Dr. Sander Greenland, "analysis of heterogeneity can be the most important function of a meta-analysis." (Rothman et al. Modern Epidemiology 2008 p.671) and analysis of heterogeneity is "often more important than computing a fictional common or 'average' effect."

(Rothman et al. Modern Epidemiology 2008 p.671) Importantly, Dr. Greenland has written, "synthetic meta-analyses that ignore heterogeneity should indeed be banned from publication, if only because they violate sound statistical and scientific practice." (Greenland, Am. J. Epidemiol 1994). In my opinion, discussed in detail below, FDA failed to properly investigate and account for heterogeneity of the study results of the 199 clinical trials in its AED meta-analysis.

The FDA study had a number of design features and flaws which tend to bias the results toward finding a positive association between the group of 11 antiepileptic drugs and suicidality. These are explained in detail below.

A. FDA's methods exclude most of the data

The majority of studies submitted to the FDA, including most (at least 89.3%) of the trials for Neurontin, were excluded from the main analysis because there were no events. This would enrich the analytic cohort with three drugs which were known to be associated with an increased risk of suicide attempt based on randomized clinical trials, probably the very same trials which were included in this meta-analysis. As a result, the summary odds ratio is biased toward a positive association.

> *"The primary analysis method was the exact method for a stratified odds ratio and associated 95% confidence interval (Cytel 2005, Ch. 19). ....a large number of the trials were expected to have no events. The exact method for a stratified odds ratio does not make use of these trials."* (Levenson M et al. 2008 at Page 14.)

It is not possible to determine exactly how large this bias is based on the information provided in the FDA report. The report does not state the number of trials, overall and by drug, which were ultimately included in the calculation of odds ratios. Similarly, the number of patients, overall and by drug, which contributed to the odds ratio calculations are also missing in the FDA report. The totals provided within figures in the FDA report are misleading as they are the total of all patients as received from the companies and not as analyzed.

An inherent bias in FDA's analysis of AEDs is its use of a methodology that involves excluding trials with zero suicidality events. The net effect of this methodology is that the final result is driven by the results of studies of 3 drugs; two of which were positively associated with suicidality (Topiramate and Lamotrigine) and one that appeared to be protective (Divalproex). (Levenson M et al. 2008, Page 27) Although 199 clinical trials were submitted to FDA, with a total of 142 suicidal behavior and ideation (treatment and placebo arms combined) (Levenson M et al. 2008 Table 13, at page 23), given the rarity of suicidality events, there were likely a substantial number of trials with no events. (Katz R. 2008) In an attempt to estimate the number of trials which were excluded from the calculation of the summary odds ratio, I estimated the minimum number of trials that had to be excluded for those drugs which had five or fewer events. The majority of drugs (n=6, 54.5%) had between 0 and 5 events. (carbamazepine, felbamate, gabapentin,

oxcarbazene, tiagabine, zonisamide) (Levenson M et al. 2008 Table 13, at page 23) Assuming just one event occurred in each trial, up to the maximum number of trials submitted (Levenson M et al. 2008, Table 4 at page 11); these six drugs could have contributed no more than 15 clinical trials to the odds ratio calculation. This means that three quarters (74.1%) or more of the 58 submitted trials for these 6 drugs were excluded from the main and subgroup analyses. Expanding this estimate to include those agents that had up to 10 events, this adds pregabalin (9 events) and levetiracetam (10 events). (Levenson M et al. 2008, Table 13 at page 23) This brings the total drugs to 8 of the 11 in the FDA's analysis. These 8 agents contributed a total of 117 placebo controlled clinical trials to the FDA for analysis, though a maximum of 34 trials (29.1% of those submitted) would have contributed to the main analyses. This illustrates the inherent bias of excluding zero event trials when zero event trails are common among the eligible trials. With the two drugs, lamotrigine and tomopiramate, there were more suicidal events than trials. Thus, most or all of the trials may have been retained in the analysis. This means that the FDA's summary odds ratio is weighted – to a large extent – by the results of two drugs (lamotragine and topiramate), which were labeled for suicide attempt prior to the study being designed and conducted. These two drugs appear to drive the odds ratio in a positive direction.

B. FDA did not use mutually exclusive categories for classification of the AEDs.

When FDA "categorized" drugs into three groups based on what is believed to be their main mechanism of action, they put the drug topiramate and zonisamide, both of which were shown in premarketing clinical trials to have an increased risk of suicide attempts, into all three and two groups, respectively. (Levenson M et al. 2008, page 13)

> *"Note that the drug groups are not mutually exclusive or exhaustive."* (Levenson M et al. 2008, page 13)

This is unorthodox and not generally excepted method of categorization. When categorizing a variable into groups, the groups should be mutually exclusive and exhaustive. That means that each antiepileptic drug should fit into one group and only one group. But in the FDA study the groups were not mutually exclusive. Groups 1 and 3 contained two agents, and group 2 one agent, which were known to have statistically elevated rates of the suicide attempt (and were so labeled) prior to the FDA's study protocol being developed. (ORTHO-McNeil 2006, UCB, Inc 2006, Eisai, 2007) This flawed classification scheme was further compromised by the use of the analytic methods which further enriched each group with the trials of those very same drugs.

> *"Topiramate is in all three drug groups, and there were a large number of patients from topiramate trials. Therefore, treatment effects of topiramate can be expected to have large influence on all three drug classes."* (Levenson M et al. 2008, Figure 7, page 32.)

Given that this design choice enriched all of these groups with drugs that were known to be associated with suicide attempt, it is hardly surprising that they identified an elevated rate of events in each of these "drug groups." Instead of providing evidence of consistency across different mechanisms of action, this analysis provides no additional information and instead is misleading.

C. Trial indication groups are not mutually exclusive

Another subgroup analysis that the FDA relied upon was classification and comparison of results across "Trial Indication." (Levenson M et al. 2008, Table 3, page 10 and page 13) Trial indications were collapsed by Dr. Mentari into 21 groups and then further collapsed into three categories: Epilepsy, Psychiatric, or Other. Just as topiramate was included in all three of the "drug class" groupings, it is labeled for both epilepsy and migraine indications (ORTHO-McNeil 2006), and would have contributed trials to at least two (Epilepsy and Other) of these categories. Therefore, this trial indication subgroup analysis has many of the same limitations as stated above (part B).

Available information on the distribution of patients in each of these Trial Indication Groups highlights the heterogeneity of the drugs included in the meta-analysis. While all of these drugs were selected for analysis because they had at least one epilepsy-related indication most patients were in trials for indications other than epilepsy. (Levenson M et al 2008, Table 8, page 18) The epilepsy indication group included only one-quarter of the total patients (24.4% treatment and 25.9% placebo) and an even smaller proportion (22.1% of events that occurred on treatment and 10.5% of events that occurred on placebo) of total events. (Levenson M et al. 2008, Figure 8, page 33)

As trials with no events are excluded from the main analysis, these categories become saturated with drugs, and trials from those drugs, which were known to have elevated risk of suicide attempts before the study began (i.e., lamotragine and topiramate). This biases the results toward finding a significant elevation in risk. As with the mechanism-based groupings analysis, estimates of the risk difference (which did not exclude zero event trials) were not reported.

Because the FDA did not provide a breakdown of this classification scheme by therapeutic agent, trials, and patient and events per trial, all of which are critical to interpreting the meaning and validity of these results, it is not possible to estimate the extent of such bias.

D. FDA subanalyses have critical limitations

The FDA conducted a number of subanalyses; Risk difference, Trial Heterogeneity, and Duration Difference. (Levenson M et al 2008, pages 14-15) Each of these has limitations which impact their validity. It is important to consider the extent and potential impact of each limitation, as the FDA relies heavily on these "sensitivity analyses" in their interpretation of the results.

D1. Risk difference

The first "sensitivity analysis" tried to examine the potential biases of excluding all of the many trials with no events from the calculation of the summary odds ratio. They used a method developed by Greenland and Robins to calculate a summary risk difference. (Levenson M et al 2008, page 14).

In the results they state:

> *"The overall risk difference was 1.79 (95% CI: 0.70, 2.87) per 1000 patients. The risk difference was greater than 0 and the confidence interval did not contain the value of 0. As was the case, for the odds ratio analysis, this result supports the finding that the drugs were associated with statistically significant increased risk of Suicidal Behavior or Ideation events relative to placebo."* (Levenson M et al 2008, page 26)

And they continue:

> *"For each of the 11 drugs, the risk difference estimate had the same direction as the odds ratio estimate relative to the null value of no effect."* (Levenson M et al 2008, page 26)

As stated in the FDA report (Levenson M et al. 2008, page 26), the drug-specific risk differences each went in the same direction as the respective drug-specific odds ratios. But what the FDA report neglected to mention was that, as with the odds ratios, there was also no consistency in the risk differences across drugs. Based on the point estimates, two drugs appeared to be protective (carbamazepine and divalproex), two drugs appeared to have no effect (gabapentin and pregabalin), and one drug had no events (felbamate). In summary, about one-half of the 11 drugs studied had results that were inconsistent with the summary estimate of a risk difference. No quantitative testing was conducted to estimate heterogeneity of the risk difference across trials or across drugs.

No sensitivity analysis was provided to determine if there are individual trials, or most importantly individual drugs, which had a large influence on the overall risk difference. Based on a qualitative assessment of the data provided, the summary risk difference appears to be driven by two drugs with statistically significant risk differences, lamotrigine (RD=5.40, 95% CI: 0.24-10.57) and topiramate (RD=3.05, 95% CI: 0.98-5.11). (Levenson M et al 2008, Figure 4, page 26) These two drugs provide 38.1% of the exposed patients and 64.4% of the events among those exposed.

D2. Trial Heterogenity of Effect

The results of a meta-analysis, as conducted here, is a single summary odds ratio. That odds ratio is an (weighted) average of the odds ratios of each of the included trials. Before one accepts that odds ratio as a valid summary of results across the trials it is

necessary to determine if the results are consistent across trials. (Thompson SG. 1994) The FDA used the Zelen test to estimate heterogeneity.

> *"Zelen's test....was used to test the hypothesis of a common odds ratio. However, because of the small number of events, it was expected that there would be little power to detect heterogeneity of the odds ratio across trials. The result of the test was intended for qualitative purposes."* (Levenson M et al. 2008. page 15.)

Thus, in developing the study protocol, FDA recognized that they would not be able to perform a valid assessment of trial heterogeneity. Since any statistical test is by nature quantitative, how they were going to use the results "qualitatively" is unclear. They present a p-value for the Zelen test in the results, and note that it provides no evidence for, nor does it rule out, trial heterogeneity. The low statistical power of such tests may result in the inability to detect even "moderate" heterogeneity. (Thompson SG. 1994)

> *"Fixed effect methods assume that all the trials had a common treatment effect. The p-value based on Zelen's test for the null hypothesis that all trials had a common odds ratio test was 0.735. This value does not provide evidence for trial heterogeneity in the odds ratio. However, the lack evidence does not imply that there was no trial heterogeneity."*(Levenson M et al 2008. page 26)

The test for trial heterogeneity excludes trials for which there are no events by giving them a weight of zero.

> *" The trial weight of the Mantel-Haenszel odds ratio estimator was used to quantitatively identify trials with large influence. The weight was equal to (control patients with events)*(test patients without events)/(total patients). Trials with no events had a weight of zero."* (Levenson M et al. 2008. page 15)

Just as in the main analysis, the exclusion of nonevent trials eliminates three-quarters or more of the submitted trials for 6 of the 11 drugs studied. This enriches the heterogeneity analysis dataset in the same way as it did the main analysis dataset, towards containing trials of those drugs which were already associated with suicidality.

Another example of flawed logic is the reliance on an analysis to identify individual trials with large influence (Levenson M et al. 2008, page 26-27), and using that as further evidence that the results are consistent across drugs without actually determining the influence of any single agent on the summary results. This study applied meta-analysis techniques to combine the results of hundreds of clinical trials *across 11 different therapeutic agents*. Therefore, it is critically important to determine if there are *individual drugs* with a large influence on the calculation of a summary odds ratios and/or risk differences. Such an analysis was not conducted. When the one trial identified as having the largest impact on the summary odds ratio was excluded, the summary odds ratio jumped from to 1.80 to 2.12. The trial with the most influence was with the study drug Divalproex (Levenson M et al. 2008. page 27), which in the drug-

specific analysis appeared protective with an odds ratio of 0.72 (0.29-1.84). (Levenson M et al. 2008. Figure 2, page 24)

When the FDA alert and report states that the results are "generally consistent" across the drugs studied, the authors ignore the considerable heterogeneity (variation) in the odds ratios across drugs, ranging from negative to positive. (Levenson M et al. 2008. page 6 and FDA Jan 31, 2008) They make this statement despite the clear qualitative evidence (i.e. the agent-specific odds ratios go in different directions) and without conducting a valid test of the heterogeneity of the odds ratio across agents. This is important as the FDA later relies heavily on the assumption that the odds ratios are not heterogeneous across trial to say that this signal of an increased risk of suicidality applies to all the 11 drugs studied.

Not only was there considerable heterogeneity in effect across the individual agents, there was also heterogeneity in the effect in many of the subgroup analyses including the trial indication, patient age and gender, location of the trial, and whether the comparator was placebo or low-dose. (Levenson M et al. 2008, pages 35-42) In the age-group analysis, the odds ratios ranged from 0.82 (aged 25-30 years) to 4.26 (aged 5-17 years). (Levenson M et al. 2008, Figure 9, page 35) The summary odds ratio for trials in North America was a non-significant 1.38, compared to a statistically elevated odds ratio of 4.53 for Non-North American trials. (Levenson M et al. 2008, Figure 13, page 39). In discussing the critical importance of looking at the underlying clinical factors which may influence the heterogeneity of results in a meta-analysis, Dr. Thompson states: "Meta-analysis ignoring these factors may well be misleading." (Thompson SG. 1994)

Because of the presence of considerable heterogeneity across the trials examined in this meta-analysis, it is incorrect to conclude, based on FDA's analysis, that Neurontin, itself, has an increased risk for suicidality. As such, one cannot infer from FDA's analysis that Neurontin causes or is associated with suicidality.

D3. Duration Difference

Proportional hazard's models, which estimate the relative hazard of an event, are valuable to analyze study cohorts that are dynamic, where patients may be under observation for various lengths of time and may leave without having the event (censored). It appears from the methods that the FDA used a proportion hazards model.

> *"The method was similar to the primary method, but used person-time rather than patients as the unit of analysis (Cytel 2005, Ch. 15)"* (Levenson M et al. 2008, page 15)

Meeting the assumption of constant hazards is critical to calculating a valid measure of the relative hazard between two groups. If the hazard is not constant over the observation period, then a single summary estimate of this hazard (the relative hazard) is misleading.

> *The FDA report states "Because the duration difference was small, an assumption of constant hazards was not key."* (Levenson M et al. 2008, Page 15)

This makes little sense as the assumption of constant hazards can be met (or not met) regardless of the average duration of exposure among treatment compared to placebo groups. Therefore, it is not possible to determine if the assumption of a constant hazard is met just by comparing the average duration of exposure between two groups. It is necessary to compare survival curves between exposed and unexposed to determine if the ratio of the incidence or hazard remains constant over time.

The available data suggest that the assumption of a constant hazard was not met and the person-time analysis may, therefore, be invalid. A Kaplan-Meier plot of the cumulative incidence was provided in Figure 6 (Levenson M et al. 2008, page 30) and the underlying hazards for drug and placebo are given for discrete time periods in Table 14 (Levenson M et al. 2008, page 29) of the FDA's report. The ratio of the drug to placebo hazard was not constant. Instead it appeared to increase over time, from with a low of 1.16 in the < 1 week exposure group and a high of 2.5 in the ≥ 24 week exposure group.

The unusual decision to use the most critical event (Levenson M et al. 2008, page 9) instead of the first event when looking at the "time pattern" for suicide behavior or ideation events" was not explained. For this type of analysis the first event should be used.

The results of this analysis are discussed in Section 6.3.3 and presented in Figure 5. (Levenson M et al. 2008, page 28) The title for Figure 5 is as follows:

> *"Figure 5: Suicidal Behavior or Ideation Rate Ratio Estimates, Placebo-Controlled Trials. Note: 18 Patients with missing or zero duration were not included in this analysis."*

However, the data presented in Figure 5 are labeled Odds Ratios and not Rate Ratios (or Hazard Ratios) which would have been calculated if they used Person-time analyses. Furthermore, as the method is "similar to the primary method" it most likely excludes the trials with zero events and thus introduces the very same biases described earlier in this report. These issues cast doubt onto the validity of the FDA's hazard analysis and subsequently, the hazard ratios for the drugs under study.

E. Ascertainment bias was not accounted for in FDA's meta-analysis.

The FDA study was not a prospectively conducted randomized clinical trial designed to test the hypothesis of an increased risk of suicidality with exposure to these antiepileptic agents. The FDA asked the sponsors to go back into previously conducted randomized clinical trials and search their databases for adverse events and terms in the narrative which might indicate suicide behaviors and suicidal thoughts (ideation). Unlike the outcomes in a randomized clinical trial, these events were not collected prospectively according to a predetermined protocol nor were they validated based on any clinical assessments. Therefore, a meta-analysis does not provide the same level of evidence as prospective, randomized, controlled clinical trials when considering issues of causality.

There is the real potential for ascertainment bias because of the methods used to identify suicidality events in FDA's analysis. Clinical trials are designed to minimize and monitor for potential biases in assessing the study outcome. But suicidality was not a study outcome in any of these clinical trials and therefore, the studies would not have had a protocol to prospectively monitor, identify, and validate such events. (Levenson M et al. 2008. page 8) The events, particularly suicidal ideation and behaviors, are more likely to be identified among those on the active treatment than on placebo, because patients on treatment are more likely to have and seek medical care for other treatment-related side effects than patients on placebo. (Posner et al. 2007) Because most agents had very few events, misclassification bias even if very minor, could change the findings of the entire meta-analysis. For example, both gabapentin and oxcarbazepine both had two events among exposed and one event among unexposed. (Levenson M et al. 2008, Table 13, page 23). For either of these two drugs, if just one of the exposed cases was the result of this bias, the effect is so great that it completely reverses the direction of the odds ratio. Eliminating just one event for either of these drugs would result in a crude odds ratio of 0.70 for gabapentin or 0.62 for oxcarbazepine, which would mean a protective effect.

Based on the weaknesses of the FDA study described above, there are no bases to conclude that the FDA meta-analysis shows a signal for Neurontin for suicidality.

Opinion 2. The FDA study does not corroborate the Plaintiffs' experts' assertion of a causal link between Neurontin and suicidality.

The Plantiff's assertion that the meta-analysis employed by the FDA to test for a statistical association between a group of drugs with antiepileptic indications and suicidality somehow provides causal evidence for one specific agent is without scientific validity. They use this argument to support their assertion that gabapentin causes suicide. This is without merit. First, the FDA did not conduct a clinical trial. They did a meta-analysis of almost 200 previously conducted randomized clinical trials for 11 agents all with at least one indication for epilepsy. The outcome of interest, suicidality, was a constellation of clinical terms ranging from thoughts (suicidal ideation) to behaviors (suicide attempt, completed suicide, etc.). These events were collected as adverse events and were not identified as study outcomes prior to the conduct of the clinical trials. There were no specific protocol for their collection, nor were they validated in any way. Therefore, although FDA's meta-analysis examined only randomized clinical trials, the results of the meta-analysis do not provide the same level scientific reliability as a randomized clinical trial. Randomized clinical trials and meta-analyses of clinical trial results are different; each has a different purpose and each has different inferences that can be made from the results. A meta-analysis is merely a tool for synthesizing results of clinical trials. As plaintiffs' expert, Dr. Greenland has written, "meta-analyses run the risk of appearing to give results more precise and conclusive than warranted." (Rothman Greenland & Lash 2008, page 677)

The application of meta-analysis techniques to the study of rare events, such as the FDA conducted here, is not generally accepted. Because the statistical analyses for calculated summary statistics in meta-analysis are based on large sample approximations, they may not be valid for rare (<1%) events. (Bradburn MJ et al 2007) The rate of suicidality was well below 1% in both treatment and placebo arms in the FDA's meta-analysis. (Levenson M et al. 2008, page 5) In their study, which compared the commonly available statistics under differing conditions, Bradburn et al found that the exact method (which the FDA used in their analysis) produced biased estimates for event rates <1% . (Bradburn et al. 2007) Similarly, they report that the Mantel-Haenszel risk difference method (also used by the FDA) "tended to overestimate treatment effects." (Bradburn et al. 2007, page 65) The authors summarize their findings:

> *"In summary, many of the commonly used methods for meta-analysis give inappropriate answers when data are sparse....No method gives completely unbiased estimates in any circumstances when events are rare."* (Bradburn et al. 2007, page 75)

Given the many methodological limitations and the conflicting results among the agents studied, the FDA meta-analysis does not prove a causal relationship between (the 11) antiepileptic drugs and suicidality. The FDA's uncertainty in interpreting the study results, particularly how to explain the range of effects from positive to negative, was clearly evident in the questions that they posed to the FDA Advisory Committee at a joint meeting of the Peripheral and Central Nervous Systems (PCNS) and Psychiatric Drugs (PD) Advisory Committees (PDAC)Advisory Committees held on June 10, 2008. The first two questions posed to the committee were:

1. *Does the Committee agree with the Agency's overall finding of an increase in suicidality for the 11 AEDs analyzed?; and*
2. *"Does the Committee agree with the Agency's conclusion that the finding of increased suicidality should apply to all drugs included in the analyses, despite the observation that the estimate of the Odds Ratio for 3 of the drugs was below 1? If not, to which drugs should the conclusion apply?"* (FDA Questions. 2008)

During the meeting, the lack of data was clearly noted by the FDA representative, Dr. Katz.

> *"...there are a lot of questions that we are asking you that we don't really have specific data to answer, but we do need an answer."* (FDA Transcript 2008, Dr. Katz. Page 318)

The advisory committee members made it very clear in their deliberations that they were answering the question "Is there a signal for antiepileptic drugs?" and not "do the agents studied cause suicidality?"

"I think we know the question that is before us based upon the data that we have available, do we think that there is a signal." (FDA Transcript 2008, Dr. Goldstein, Page 286)

"the Agency has concluded that there is a signal for increased suicidality for the class of AEDs" (Katz R. 2008. page 4)

"The real crux of the question is do we think that there is a signal here, is there a signal." (FDA Transcript 2008, Dr. Goldstein, page 267)

"Randomized trials do not prove causality" (FDA Transcript 2008, Dr. Hennesy, page 279)

"I vote Yes, if the question were worded that there is a statistical significant increase in suicidality, but I cannot vote Yes if the issue is causality." (FDA Transcript 2008, Dr. Rizzo, page 288)

The FDA Advisory Committee answer to the question of whether or not to apply the summary results to all of the 11 study drugs, was, based upon my experience, a conservative regulatory decision. The committee chair modified the question (FDA Questions 2008. FDA Transcript 2008, Goldstein, Page 286) so the committee could vote on a "signal" and not a causal association, as many members noted that the data was not adequate to support any causal inference.

The results were not consistent across the individual agents in the study. This was clearly stated in the FDA briefing materials and a major point of discussion among the committee members and the FDA representatives during the meeting.

"1) the finding was not seen for all drugs, and 2) given the disparate pharmacologies of the drugs, there is no obvious explanation for why there should be a common finding of an increase in suicidality." (Katz R. 2008)

" It is about what to do with the diversity of responses in a situation like this." (FDA Transcript 2008, Dr. Temple, page 291)

"For safety concerns, the conservative thing to do is to assume that all drugs of a particular class....share the same risk unless proven otherwise, that is a reasonable assumption to make in a safety context." "But I think we should be clear that we are making an assumption and that we are not able to use scientific reasoning to make the inference that each one of these individual agents has an increased risk" (FDA Transcript 2008, Dr. Hennessy, page 291-2, seconded by Dr Pine, page 292)

In their discussion, the committee members considered that, based on the agent-specific results, some agents in the study may actually be protective. However, proving the absence of risk for rare events that occur in the populations that use the agents in the

FDA meta-analysis is beyond the current capabilities of epidemiology.

> *"Trying to prove a negative with very rare events is virtually impossible..."* (FDA Transcript 2008, Dr. Goldstein, page 289)

The FDA made an assumption that these 11 agents formed a group of drugs based on the fact that they all had at least one indication for epilepsy. The record reveals that this design decision was subject to considerable debate and members of the committee seemed to articulate different rationales and concerns during this discussion.

> *"We don't know of a mechanistic explanation for this, but we think the conclusions should apply to all 11 of them (antiepileptic drugs studied) even though some of them didn't actually lean that way."* (FDA Transcript 2008, Dr. Temple, page 310)

> *"...we have a number of medicines that have different biochemical mechanisms of action and yet, are antiepileptic drugs."* (FDA Transcript 2008, Dr. Gilman, page 314)

> Dr. Twyman asks Dr. Katz "*...could you clarify what is this class definition?"* To which Dr. Katz *replies "Anything that has an indication for epilepsy, any seizure type."* And Dr Twyman notes *"...there is absolutely no data particularly for the compounds that are specific to generalized seizures."* (FDA Transcript 2008, Dr. Twyman, Page 304)

FDA's assertion that these 11 drugs should be grouped together as a class is not supported. These drugs have a variety of possible mechanisms of action and a number of different indications and uses. The FDA did not adequately justify its decision to lump such disparate drugs into an analytic group.

Committee members during the meeting expressed their position that any labeling or alert that is incorporated into the labeling of the AEDs in the future should be worded so it was clear that the data were not adequate to make specific statements about any of the individual agents in the study.

> *"...alerting the field to the paucity of data, so really thinking very carefully about how exactly limited the data are in terms of saying anything specific about any of the 11 medications would be really important."* (FDA Transcript 2008, Dr. Pine page 292-3)

> *"... I think that the degree of uncertainty around the information that we think we know ought to be present as well."* (FDA Transcript 2008, Dr. Hennessy page 347)

> *"...there is also a lack of data for a lot of this."* (FDA Transcript 2008, Dr. Goldstein, page 347)

Dr. Jackie Ware, in the Division of Neurological Products stated that her preliminary list of antiepileptic drugs had 43 unique generic names. (FDA Transcript 2008, Ware J. Page 306) The discussion of class labeling revolved around a concern that a black box warning or other labeling of just the 11 drugs included in this FDA study would shift prescribing to other antiepileptic drugs which may have an increased risk of suicidality but were not studied, and away from agents in the study, which may be safe or even protective. They expressed appropriate concern that more harm might be done by labeling the study drugs for suicidality as they may cause patients to use less safe medications.

> *"That's the overarching concern that we can shift prescribing to other drugs that might very well have the same signal."* (FDA Transcript 2008, Dr. Katz, page 298)

> *"What would clinicians do if you only had the box warning on the current generation of antiepileptics, would they be inclined to use older drugs....for which you may not have as much safety data?"* (FDA Transcript 2008, Dr. Laughner, page 303)

Ultimately, the recommendations of the committee appeared to be based on regulatory and/or policy issues concerning the broad grouping of AEDs. Examples of the regulatory nature of the decision, as expressed during the advisory committee meeting, are below:

> *" ...this is more of a policy question. What does one assume in the absence of data rather than a scientific question what do we know about the effects of drugs that weren't studied. Obviously, we don't know about the effects of drugs that aren't studied. The question is from a public policy perspective what do we do about that."* (FDA Transcript 2008, Dr. Hennessy page 299)

> *"...earlier we assumed that the results should apply even to the drugs for which there wasn't a point estimate greater than 1. So, it seems easier to me to extrapolate to areas where we don't have data to areas where we have data that seemed to suggest safety."* (FDA Transcript 2008, Dr. Hennessy page 311)

> *"..all of these conditions are associated with depression and potentially increased suicidality. So having the patient population aware of this alone, regardless of whether it's drug-related or not is an important thing to do."* (FDA Transcript 2008, Dr. Goldstein, page 349)

The recommendations of the FDA Advisory Committee are not consistent with, nor do they imply, that the FDA's meta-analysis provided any proof of causality for the group of studied drugs, individual agents within that group, or what was defined by the FDA as "antiepileptics." I think it important to address the inconsistent use of the word

"causality" by the FDA officials at the advisory committee meeting on July 11, 2008. As noted above, many committee members were quite clear and consistent in their opinion that a causal relationship between AEDs and suicidality had not been established. The FDA members were not so clear or consistent at the meeting. Being prodded by Dr. Ruth Day about his "comfort" with "saying causality" within the context of patient communication, Dr. Katz says:

> *"We are unequivocally comfortable with using the word, the "c" word with saying that this establishes causality."* (FDA Transcript 2008, Dr. Katz, page 285)

He then continues:

> *"This is how we determine causality, this is how we base our findings of effectiveness for drugs. We do randomized trials, we analyze them prospectively, we have an outcome measure, and if it's statistically significantly different from placebo, we say the drug caused it, you know, once you rule out chance and fraud and bias and that sort of thing, which we think we have done here. So, yes, we are quite comfortable with saying there is causality."(* FDA Transcript 2008, Katz, page 285)

It is seems clear from the second part of his statement, that Dr. Katz mistakenly thought that he was discussing a randomized clinical trial. His statements as reflected in the transcript of the proceedings are inherently contradictory and incorrect. The FDA's meta-analysis is not the equivalent of a prospective, randomized controlled, clinical trial. As stated by Berlin and Kim, "Meta-analysis may be regarded as a "state-of-the art" literature review." (Berlin JA and Kim CJ 2005, Chapter 44) Meta-analyses lack most of the very features that Dr. Katz stated are the hallmark of a randomized controlled clinical trial. In fact, it is beyond dispute that the design of the FDA's meta-analysis is not a randomized controlled clinical trial and does not provide nearly the same level of evidence that an efficacy trial would provide for predetermined and prospectively gathered endpoints. The meta-analysis provides no basis to make causality statements concerning any individual agent, or even the "group" as a whole.

The Chair of the FDA Advisory Committee clearly articulated that the meta-analysis under discussion did not have the features that Dr. Katz described above.

> *"...these trials were not prospectively randomized to test this question. So this is an observational analysis that is nested in a series of randomized controlled trials that we are analyzing."* (FDA Transcript 2008. Dr. Goldman page 275)

Dr. Katz was far more circumspect in his memorandum, dated June 12, 2008, which he sent to the FDA Advisory Committee members with the final report.

> *"Based on these data, the Agency has concluded that there is a signal for increased suicidality for the class of AEDs..."* (Katz R. 2008 page 4)

Dr. Katz explains in this memo that the FDA was presenting this analysis to the joint Advisory Committee to get their feedback on whether they would apply the conclusions (of a signal for an increased risk of suicidality) to "any, only some, or all" antiepileptic drugs (studied or used chronically). This uncertainty in the meaning of the analysis expressed by Dr. Katz is in stark contrast to the plaintiffs' experts' interpretation of the FDA's stance that there is evidence of a causal relationship between the 11 AEDs and suicidality and Neurontin specifically. (Plaintiffs' Opposition to Defendant's Daubert brief 2008 page 25-26.; Daubert hearing transcript at 567-68). It was clear from the advisory committee members' discussions and votes, that the committee members were cautious to not over-interpret the results of this meta-analysis.

Interestingly, although the FDA Advisory Committee voted that it agreed with FDA's findings, based upon the meta-analysis that there was a signal for increased suicidality for the group of 11 AEDs analyzed, and that this finding should apply to all currently approved AEDs, the FDA Advisory Committee voted against recommending a Black Box warning for suicidality in the labeling of AEDs (4 vs. 14 with 3 abstentions). The committee specifically rejected a Black Box warning and did not vote to recommend any other warning pertaining to suicidality. Thus, if the FDA Advisory Committee was not convinced of the need for a Black Box warning with clinical trial data from 199 studies from 11 different AEDs, there is absolutely no evidence to suggest that Pfizer, on its own, should have made changes to the Neurontin labeling as suggested by plaintiffs' experts.

It is my professional opinion that the FDA study of 11 anti-epileptic drugs did not provide proof of a causal relationship between anti-epileptic drugs as a group, or any of the individual antiepileptic drugs studied including Neurontin, and suicidality.

Opinion 3. It is my opinion that the plaintiffs' analytic methods are inappropriate for postmarketing safety analysis. There was no reason for Pfizer to conduct similar analyses to those advanced by plaintiffs as part of its postmarketing safety analyses.

A. There is no evidence that FDA approved or vetted plaintiffs' methods
In both Dr. Blume's and Mr.Altman's declarations, they imply that Mr. Altman's methods are somehow vetted and approved by the FDA because they approved two NDA's and are reviewing a third in which Mr. Altman had authored or performed analysis for a section. They state that FDA has neither rejected "my or Mr. Altman's submissions for errors, omissions, inaccuracies, or methods in his computational analysis." (Blume C. 2008, para 10) This misstates the FDA process in reviewing applications. Based on my service on FDA Advisory Committees and reading of FDA safety evaluations, the FDA performs its own independent evaluation and analysis of the safety data provided in an NDA. I am not aware of the FDA's grading all or parts of NDA's like an exam nor summarily rejecting an NDA based on errors, omissions, or inappropriate methods within this one section. Nor is there any evidence that the FDA accepted their methodology. I am not aware that Dr. Blume has even disclosed the specific of her methods used in her regulatory submissions or any FDA response thereto.

B. Dr. Blume misstates the contents of 21 CFR 201.57(e)

Dr. Blume misinterprets the federal regulations regarding the level of evidence required for a black box warning (21 CFR 201.57(e)) to justify why her work should be accepted even though she did not use scientifically valid methodology. (Blume C. Declaration. 2008 para 19)

> *"It is not universally necessary to employ the various methods of epidemiology to establish whether there is an association between a drug and a risk. For example, the FDA in 21 C.F.R. 201.57(e) states as such that epidemiology is not indispensable to providing even a black box warning."* (Blume C. Declaration. 2008, para 19)

Contrary to her declaration, the section which she references in the C.F.R. has nothing to do with epidemiological methods. It only states that *human data* shall be in most cases, but not always, the basis for a boxed warning. This gives the FDA the regulatory authority to require a boxed warning based on animal data alone, when human data are not available. Thus, in some situations the FDA may decide to warn about a particular risk identified in animal data, even if there is no evidence that it is applicable to humans, when they believe that such a warning will serve the public health.

> *"Special problems, particularly those that may lead to death or serious injury, may be required by the Food and Drug Administration to be placed in a prominently displayed box. The boxed warning ordinarily shall be based on clinical data, but serious animal toxicity may also be the basis of a boxed warning in the absence of clinical data."* (21 C.F.R. 201.57)

C. Plaintiffs' methods are not approved or generally accepted for postmarketing safety analysis or causal assessment.

Dr. Blume further justifies her "methods", particularly her production of numerous tables with line listings of events and the lack of statistical testing by saying that these are the same type of tables she creates as part of a new drug application (NDA). The purpose of an NDA is to present regulators with the evidence that the drug's benefits outweigh its risks. The use of summary tables in regulatory submissions is to facilitate FDA's review of the relevant data. Such presentations do not, in and of themselves, permit a reliable analysis of causation. There is ample evidence in this case that FDA would only consider randomized clinical trial data to assess whether AEDs are associated with suicidality.

The purpose of a pharmacoepidemiology study is to test a hypothesis, often to determine if a drug is associated with an increased risk of a particular adverse outcome or outcomes. These very different purposes call for very different methods. Hypothesis testing requires appropriate statistical analysis.

Dr. Blume consistently argues that her expertise in drug development and her use of the same methods that she uses to prepare tables for a NDA are appropriate to test postmarketing safety hypothesis.

> *"My professional life has been directed to the development of pharmaceutical projects for submission to the United States Food and Drug Administration on behalf of pharmaceutical manufacturers."* (Blume Declaration 2008, Para 3)

> *"...I used the same methods which I used in performing a drug development submission to the FDA." (Blume Declaration 2008, para 5)*

> *"These records...are the type of records that the FDA routinely reviews in its evaluation a pharmaceutical product to evaluate risks and benefits and to evaluate safety and efficacy. I used the records the same way the FDA uses them,..."* (Blume Declaration 2008, para 9)

> *"The methodology which I employed in this case is similar to those I employed in my drug development and safety surveillance projects ... It also is the same methodology Pfizer used in NDA's in this case... "* (Blume Declaration 2008, para *15)*

> *"My charts and graphs are similar to those used by the company within its NDA's for Neurontin." (*Blume Declaration 2008, para 34)

She states that she does not use epidemiologic methods in this situation because they are not necessary.

> *"It is not universally necessary to employ the various methods of epidemiology to establish whether there is an association between a drug and a risk."* (Blume Declaration 2008, para 19)

This ignores important concepts in the practice of pharmacoepidemiology. Case report listings are not used to determine whether or not there is an association between a drug and an event. To test for an (statistical) association requires a statistical test.

> *"Case reports are useful for raising hypotheses about drug effects, to be tested with more rigorous study designs. However, in a case report one cannot know if the patient reported is either typical of those with the exposure or typical of those with the disease. Certainly, one cannot usually determine whether the adverse outcome was due to the drug exposure or would have happened anyway."* (Strom BL. 2005. Chapter 2)

In paragraphs 18 and 19, Dr. Blume makes contradictory statements concerning whether or not she used, and the importance of, background rates. (Blume C. Declaration 2008)

In paragraph 18, she states that she considered background rates (of what is not specified) of the Neurontin patient populations to provide context to clinical trial results, but in paragraph 19, she argues that using background rates to provide context to adverse event reports "is contrary to good pharmacoepidemiology practices." This is wrong. As stated in the FDA guidance document, "To provide further context for incidence rates or reporting rates, it is helpful to have an estimate of the background rate of occurrence for the event being evaluated in the general population or, ideally, in a subpopulation with characteristics similar to that of the exposed population." (FDA Good Pharmacovigilance Practices and Pharmacoepidemiology Assessment, March 2005)

In paragraph 21, Dr. Blume states that "*there is no epidemiologic evidence that a causal relationship does not exist between Neurontin and suicidality.*" (Blume Declaration 2008, para 21) This is correct, not because of the data or lack thereof, but because of the limitations of scientific research. Researchers are unable to conduct experiments that prove the absence of risk. One can only test to see if there is evidence of an increase or decrease in risk. The randomized, controlled clinical trial data do not provide evidence to suggest either an increase or decrease in the risk of suicidality or suicide with Neurontin, which is consistent with Pfizer's statements to FDA in 2004 and 2006.

Individual case reports involving patients who were dechallenged (taken off the medication) and the reaction resolved and subsequently rechallenged (put back on the medication and the reaction reoccurred) may in rare circumstances provide evidence of a causal relationship between the drug and an adverse effect. However, there are many caveats to the use of such events that are generally accepted. First, patients may get better and relapse as part of the natural history of disease. Second, other factors which also vary over time, other than the drug, may be the cause of the event. Third, subjective adverse events that are reported following a rechallenge are more prone to biased assessment compared to events which can be quantitatively assessed (e.g. blood levels). Also, conditions that tend to occur episodically, such as depression or headache may at some point cycle with drug use mimicking dechallenge/rechallenge. Fourth, a rechallenge can be inherently suggestive and therefore, biased for such subjective patient-reported events. Without accounting for these biases, one cannot assess whether a particular drug causes this type of subjective event based on a mere report of a dechallenge and/or rechallenge.. (Knowles, et al. 2002). Given that the events that Dr. Blume is using as evidence for dechallenge/rechallenge are both highly subjective and episodic, her conclusions are unjustified.

In paragraph 29, Dr. Blume compares controlled and uncontrolled clinical studies and makes light of the difference. (Blume Declaration 2008, para 29) In an uncontrolled study there is no randomization to treatment and control groups. Everyone gets the active treatment. Because there is no placebo group there is no basis to evaluate benefits and risks. This is particularly critical when evaluating the outcome (suicidal behaviors) of interest, as the background rate is known to be elevated among most of the indications for which Neurontin was evaluated in uncontrolled studies. Such limitations apply to all

antiepileptic drugs, which explains why FDA only included data from randomized placebo controlled clinical trials in its meta-analysis.

Opinion 4. Using the published and accepted methods for calculating PRR, there was no signal of disproportional reporting (SDR) for reports of completed suicide with Neurontin until 2005 and suicide attempt until 2006. This is consistent whether compared to a background of all drugs or the antiepileptic drugs in the FDA analysis. There was no SDR for Suicide attempt when compared to the other ADES to date. Based on my analysis of the evidence, Pfizer is correct in concluding that it did not see a signal for suicidality with Neurontin in the adverse event reports.

FDA maintains a database of adverse event reports which they receive from manufacturers as well as directly from healthcare professionals and the public. In November 1997, this Adverse Event Reporting Database (AERS) was significantly modified and many changes happened concurrently. These included, but are not limited to, a change in dictionaries used to code adverse events (from COSTART to MeDDRA), a lifting of limits on the number of adverse event terms that could be listed on each adverse event report, and not entering (periodic) manufacturers nonserious adverse event reports into the database. (Woodcock J. 2002) Since its introduction, there have been numerous modifications to the MedDRA dictionary. (MedDRA Release Archive) Changes in reporting behavior, trends in the use of particular adverse event terms, the period report waiver program, and a high level of publicity surrounding certain drugs and adverse events complicates the evaluation of adverse event reporting rates and renders comparisons, particularly comparisons across individual drugs, highly problematic.

Data mining, as used in the field of pharmacovigilance, is the use of statistical algorithms to identify unexpected variations in reporting rates. While it is currently being used and/or evaluated for use in a number of companies and regulatory agencies, its utility is still widely debated and methods vary substantially. (Weiss S. et al. 2008) At a recent meeting (an invitation-only think tank held on June 2008) of approximately 100 pharmacovigilance experts from the United States, which I chaired, use and value of data mining in pharmacovigilance was widely debated.

Even in 2008, data mining is not considered standard practice in the pharmaceutical industry. In the FDA Guidance to Industry on Pharmacovigilance and Pharmacoepidemiology, which was published in March 2005, it states "*Use of data mining techniques is not a required part of signal identification or evaluation. If data mining results are submitted to FDA, they should be presented in the larger appropriate clinical epidemiological context.*" (FDA. Guidance to Industry on Pharmacovigilance and Pharmacoepidemiology 2005)

There was no consensus on what constitutes a significant difference in reporting rates (a signal of disproportional reporting or SDR) nor was there agreement on acceptable protocols or best practices. Also, even when there is an alert or signal of disproportional reporting from a data mining excursion, it does not establish causality. It is generally accepted in pharmacovigilance that a finding of disproportional reporting is without any

clinical meaning. Only with careful clinical interpretation, considered within the context of why the drug is being used, the extent of drug use, and background rates of the event, can a statistical alert be turned into what is considered a signal of a potential risk. (Hauben M et al. 2005, FDA Guidance Pharmacovigilance and Pharmacoepidemiology 2005, Ahmad SR et al. 2005) A number of data mining algorithms (though not all) can be easily run on AERS data by anyone with a home computer and spreadsheet software, the interpretation of such analysis is complex. I see no evidence that Dr. Blume or Mr. Altman enlisted proper medical experts to interpret their purported signal.

Using a basic epidemiological formula for relative risk, I calculated what is called a proportional reporting rate (PRR) (Hauben M and Zhou X. 2003), comparing the proportion of gabapentin reports received by each time point that included the MedDRA term for "suicide" and "suicide attempt" to the proportion of all other AERS reports which included that term. The results are in my original report. My analysis was limited to reports received after gabapentin was approved for marketing (1994). The PRR was recalculated annually based on all the reports received by the FDA through that year.

I understand that I was criticized by Mr. Altman, in his declaration for graphing values of zero. (Altman K. Declaration 2008, para 37). Zero is the software's default for no value. (Qscan User Manual Version 3.2) PRR cannot be calculated for terms that are not reported. The baseline value for a PRR – which signifies no difference in reporting rates – is 1.0. A PRR above 1.0 means that particular term is more common (proportionally) among reports for the drug under study than among the background (typically all other) drugs. There is no consensus in the pharmacovigilance community on what PRR value is considered to be in excess of expected (statistically significant), though a value of 2.0 or above, often with additional requirements such as a chi-squared above 4 and at least 3 drug-event pairs, is frequently cited in the literature. I graphed the PRR's, not cumulative percentages, as erroneously noted by Mr. Altman. The start date was selected because that was the year gabapentin was first approved and the reporting of the first suicide attempt reports. The jump in my graph from zero to 1.0 does not mean that there was an increase in the PRR, but that the PRR could now be calculated for that preferred term.

Mr. Altman also criticizes my analysis of completed suicide because events with that term are not in the AERS database prior to 1997. (Altman K. Declaration para 37) Prior to November 1997 adverse events were coded using the COSTART dictionary and not MedDRA. When this legacy data was merged with the newer AERS data, the more general COSTART terms were mapped to MedDRA terms. The suicides and suicide attempts from COSTART were mapped to the MedDRA PT "suicide attempt." In my original report the PRR for both completed suicides and suicide attempt are plotted. Contrary to Mr. Altman's argument, given that the first two suicide attempt reports were reported in 1994, it would have been misleading to throw out these reports which were received before 1997.

In my initial analysis, I compared reporting rates for Neurontin (generic and brand) to a background of all drugs in the FDA's database. Given the FDA's meta-analysis, I analyzed the AERS database to compare Neurontin to a background of those AEDs

included in FDA's study. This would have the impact of reducing bias in the PRR based on the indication for treatment (e.g. the elevated background rates of suicide among epileptics compared to the general population). Because of the changes in FDA's database over time, and the variable length of time these drugs have been marketed, I limited the analysis to reports received on or after November 1, 1997. Each data point accumulates reports from November 1, 1997 through October 30th of the year noted in the graph. If there were no reports for Neurontin the PRR could not be calculated and is not plotted.

Based on this analysis, there was no SDR for Neurontin (gabapentin) for the term "completed suicide" until 2005. (Figure 1) Indeed, the PRR remained below one, which would suggest that there were less completed suicide reports than would be expected, through 2003. By that time, it seems that reports of suicide and suicide attempt were inflated for Neurontin in the AERS database and the increased PRR was an artifact of reporting bias. (Altman K. Declaration 2008, para 24, Blume C. Deposition 2007, page 187)

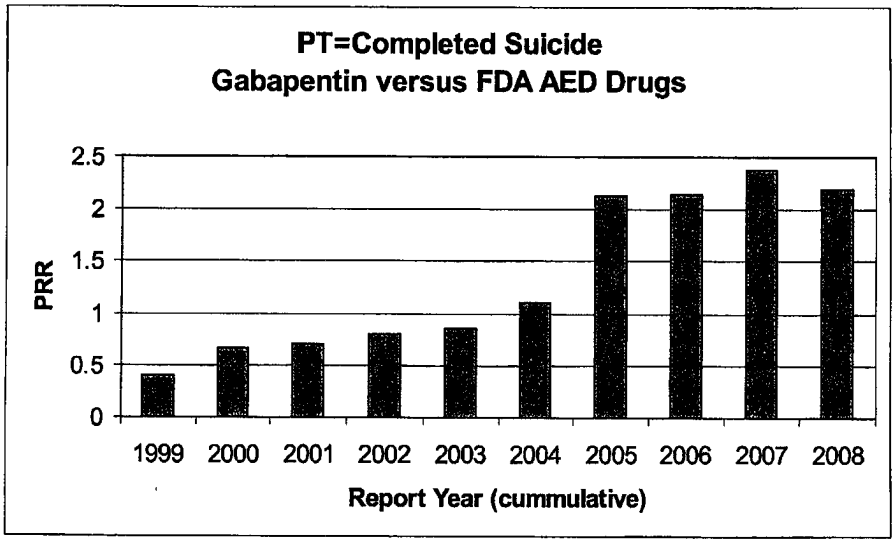


Figure 1. PRR for the term Completed Suicide with Neurontin (gabapentin) against a background of the drugs in the FDA's AED meta-analysis, cumulatively by year.

Similarly, there was no SDR for the preferred term "suicide attempt" for Neurontin when compared to the other AED drugs. (Figure 2) Again, the PRR suggested that these events were reported less than expected through 2005.

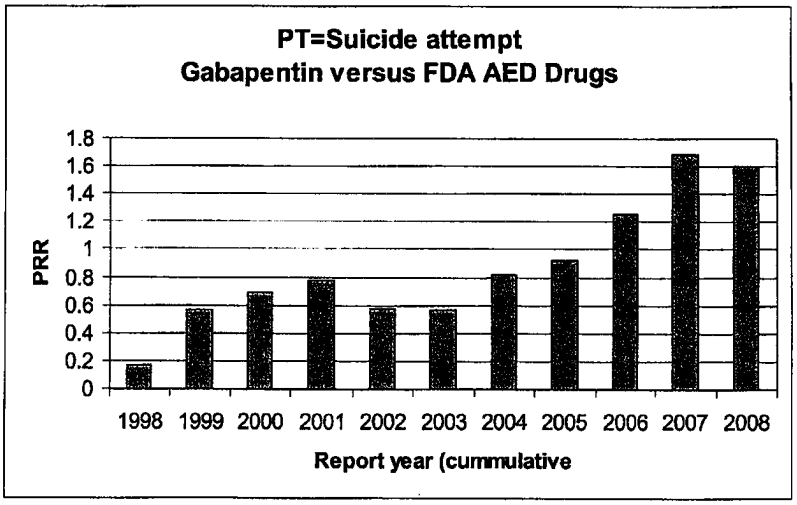


Figure 2. PRR for the term Suicide Attempt with Neurontin (gabapentin) against a background of the drugs in the FDA's AED meta-analysis, cumulatively by year.

Mr. Altman implies that I should have used the "HLT" Suicide and self-injurious behavior" which includes a constellation of terms ranging from completed suicide to self-injurious ideation. While I disagree with this, I did look at the PRR for this HLT in the same analysis. Again, the PRR was at or below 1 (which is the baseline value indicating no difference) through 2004.

Therefore, data mining showed no statistical signal for Neurontin for these events until the reporting rates were changed by an influx of Neurontin reports in 2005.

Opinion 5. Plaintiffs have not provided a "corrected" version of the data mining exercise that I had included in my original report. Rather, they present an entirely new analysis using methods that are not generally recognized as data mining, nor accepted as useful for postmarketing surveillance of adverse events.

Mr. Altman's states that he created a "corrected version" of my data mining analysis. (Altman Declaration 2008 para 38). But he did not calculate any recognized data mining statistic. Furthermore, he makes a number of substantial methodological changes such that it is an entirely new and different analysis. (Altman Declaration, 2008 para 39 sections A-F) Therefore, his analysis (Altman Declaration 2008 – Exhibit D) is unrelated to, and not a correction of, the analysis in my original report.

A. Primary Term versus Higher Level Term

The analysis I conducted was at preferred term level of the MedDRA dictionary and included two terms, completed suicide and suicide attempt. Both of these events meet FDA's regulatory definition of serious adverse events (21 C.F.R. 314.80) Instead of doing the same, Mr. Altman used a higher level of the MedDRA dictionary (HLT) 19.23.1 "suicidal and self-injurious behavior" which includes all of the following

preferred terms; completed suicide, intentional self-injury, self injurious behaviour, self-injurious ideation, suicidal behaviour, suicide ideation, and suicide attempt. The use of this broad category includes events which would meet the regulatory definition of serious as well as those which would not, such as self-injurious and suicidal ideations. (21 C.F.R. 314.80) Limiting his analysis to "serious" reports based on the patient's outcome does not correct any alleged problem with my analysis.

In her declaration (Blume C. Declaration 2008, para 32), Dr. Blume justifies the use of the MedDRA HLT "suicidal and self-injurious behavior" by saying that because they are a "standard dictionary collection" in the MedDRA dictionary that together these terms represent a clinically meaningful concept. (Blume C. Declaration 2008. para 32) This is wrong. The ICH sanctioned working group on the use of MEdDRA for data retrieval which noted that HLT's *"should be viewed as an additional tool to aid in data retrieval." They recommend that users review all terms within a HLT to "ensure that all terms are suited for the purposes of the output."* (MedDRA Term Selection, page 9) They further state that *"Clinically related PTs in MedDRA might be overlooked or not recognized as belonging together as they might exist in different locations within a single SOC or within more than one SOC,"* (MedDRA Term Selection, page 11) and therefore they recommend that an individual with a medical background, who is also trained in the use of MedDRA review the data retrieval and presentation strategy. (MedDRA Term Selection, page 11) Recognizing the complexity of grouping MedDRA preferred terms into clinically meaningful concepts, there is an international working group assigned the task of creating such groupings, which are called Standardized MedDRA Queries (SMQ's). (MedDRA Term Selection, Page 11)

Dr. Blume (Blume Declaration 2008, para 39) further justifies her use of the HLT "suicidal and self-injurious behavior" based on her observation that in the company's "Gabapentin Data Capture Aid (Blume Declaration 2008. Exhibit G) the company used all of the preferred terms under this HLT "plus a few additional terms." In this document, the company listed a total of 10 terms under the heading "potentially suicide/self-injury related adverse events." By using the HLT, Dr. Blume used only 6 (60%) of the terms in Pfizer's data capture aid. One of the terms she did not include, depression suicidal is listed under the HLT "Depressive Disorders" under the SOC Psychiatric disorders. Two additional terms which she missed, intentional overdose and poisoning deliberate, are under a different SOC, Injury Poisoning. Furthermore, the company is using this as an aid for their medical staff to identify a pool of incoming reports that required more detailed information than typically collected. (Blume Declaration 2008. Exhibit G). This is to facilitate data capture of supplemental data for events deemed of particular interest, not data analysis for the purpose of signal detection. Therefore, her assertion that the company "did the very same thing" and that the company in some way "validated" her approach is incorrect. (Blume Declaration 2008, para 39)

Reviewing the Data Capture Aid Dr. Blume states that "*there is no technical reason why such a protocol could not have been implemented in 1994 based on the signals already existing with respect to Neurontin and suicidal behavior.*" (Blume Declaration 2008, para

38) This is incorrect. Even if there were a signal in 1994, which there is no evidence to support, the Data Capture Aid could not have been done in 1994. As the plaintiffs' stated in their critique of my analysis (Blume Declaration 2008 para 37 and Altman Declaration 2008, para 37 section a), the preferred term "completed suicide" was not in the database prior to 1997. The majority of adverse event terms in the Data Capture Aid did not exist in the COSTART dictionary in 1994 (e.g. intentional self-injury, self-injurious ideation, poisoning, deliberate, self-mutilation, suicide ideation, completed suicide). (Morse R. 1994).

B. Serious and suspect

Mr. Altman purports to limit his analysis to reports with serious outcomes as part of his alleged efforts to "correct" my analysis. Serious is a regulatory definition based on patient outcome and not the reaction terms coded. Serious is defined as events that lead to death, hospitalization (initial or prolonged), life threatening, persistent or significant disability, birth defect, or an important medical event that would have resulted in any of these outcomes if not for medical or surgical intervention. (21 C.F.R. 314.80) Just because a case report meets the definition of a serious outcome, this does not mean that a suicidal event which might be listed in that report met the definition of "serious", nor that it was the event leading to the report being filed. With the advent of AERS in November 1997, an unlimited number of event terms can be entered for any one report. This led to information from the narrative, which included underlying conditions and indications for the medications, being coded as events in AERS reports. That is why I limited my analysis to two preferred terms suicide and suicide attempt, which are far less ambiguous and uniformly serious. Therefore, Mr. Altman's attempts to correct my analysis are misguided and his analysis is misleading.

Mr. Altman purports to limit his analysis to reports where Neurontin was indicated as the "suspect" drug, though it is not clear as written (Altman Declaration 2008, para 38 section d). The use of the suspect label has two meanings, not just the one that Mr. Altman mentions. When a report is received by the FDA directly from the public (health care professionals, patients, lawyers, etc.) the reporter may indicate which drug or drugs they suspect as causing or contributing to the event among all of the drugs that the patient was taking at that time. However, when a manufacturer submits a report, they are required to indicate their drug as the suspect drug, regardless of whether or not that makes sense clinically.

Mr. Altman's inclusion of a broad and diverse groups of terms, and then limiting the reports to those which are "serious" based on outcome does not exclude all but "serious" reports of suicidality as he contends. To illustrate this lack of sensitivity of limiting events to "serious" while using the broad HLT categoriy, I searched the FOI-AERS database for all Neurontin reports, for which the drug was marked as one of the suspect drugs, and reviewed reports which had the HLT 19.23.1, but not the preferred terms "completed suicide or suicide attempt". Two examples (of many) of reports that would be included in Mr. Altman's analysis based on his "rules" are described below.

Case ID 3610385 is an expedited (follow-up) case reported to FDA on 11/9/2000 by Janssen Research Foundation, Division of Johnson & Johnson of a 48 year-old male. The outcomes included hospitalization, disability, required intervention to prevent permanent impairment/damage, and other. This patient had 44 concomitant drugs listed, of which the primary suspect drug was Propulsid and Neurontin was one of seven secondary suspect drugs. There were 88 reported reaction terms for this patient, of which suicide ideation was one. The reported terms included, "myocardial infarction," "ventricular fibrillation," cardiomegaly," "chest pain," "mitral valve incompetence," and "Q-T prolonged." From the report, it cannot be determined what triggered this case to be classified as "serious," but it is unlikely that it was suicide ideation.

Case ID 4546166 is an expedited (follow-up) case reported to FDA on 12/30/2004 by Pfizer of a 45 year-old female weighing 120 pounds. The outcomes include hospitalization (initial or prolonged). This patient had 28 concomitant drugs listed, of which Neurontin was one. There were 189 reported reaction terms for this patient, of which suicide ideation was one. However, the reported terms included "road traffic accident," "deafness," "viral infection," "multiple fractures," "fibrocystic breast disease," head injury." Because this case was reported by Pfizer, by regulation, Neurontin is coded as a suspect drug for this case. Again, from the report, it cannot be determined what triggered this case to be classified as "serious," but it is unlikely that it was suicide ideation.

Therefore, because Mr. Altman's inclusion criteria for his analysis of "serious" cases lacks any specificity or sensitivity and would allow such cases as this one to be included in the analysis, any finding of a "signal" from his analysis would be erroneous and unreliable.

C. Not calculating a PRR

As discussed above, it is important to remember that percentages of adverse event reports are not interchangeable with PRR's or other data mining statistics. Data mining is used to identify particular adverse event terms which are reported more often than expected with the study drug. To do this requires one to calculate an expected value and then, using a statistical algorithm, calculate the test statistic. In my analysis, I used the PRR (with no additional constraints such as minimum case counts). While the various algorithms tend to give similar results when there are at least 20 cases (FDA Guidance Pharmacovigilance and Pharmacoepidemiology 2005, page 9), this is a very inclusive threshold.

Mr. Altman did no such analysis. Thus, in regards to Mr. Alman's graph in his Appendix D, I find that it provides no information that would establish a signal of disproportional reporting for any meaningful event. This graph is the results of a retrospectively determined process and is only one of a very large number of looks he made of the data. (Altman CD-ROM, 2007). Pfizer should not be faulted for not analyzing the data in a manner similar to Mr. Altman. The use of such imprecise and flawed methods would only serve to distract responsible parties from discovering relevant signals.

Opinion 6. Plaintiffs' interpretation of the postmarketing adverse event data is flawed and deviates from generally accepted concepts in pharmacoepidemiology.

Evaluation of spontaneous reports needs to be conducted within the context of the number of patients being treated by the drug and how it is being used. It is not appropriate to tally the number of events at a time point and not consider the population exposed. The plaintiff's experts listed counts and percentages of some adverse event terms without any comparative analyses or accounting for changes in the patterns of drug use and adverse event reporting. They state that their interpretation of the AERS data are somehow validated by the mere presence of similar tables in New Drug Applications (NDA's) that are submitted to the FDA for review. (Blume Declaration 2008 para 10)

While such descriptive statistics are often provided as context or background within NDA's and other regulatory submissions, they do not provide any evidence of an association. To argue that the mere inclusion of similar data tables within NDA's or Integrated Safety Summary reports as evidence that they are a valid method to evaluate risk of a particular adverse event is without any scientific merit. Indeed, in an evaluation of trends in the reporting of suicidal thoughts and behaviors among children and adolescents, FDA experts Drs. Mosholder and Palmer concluded that the evaluation of suicidal thoughts and behaviors reported in AERS were not "scientifically rigorous" due to many external factors that influence reporting. (Mosholder AD and Palmer 2006) Thus, raw counts and percentages of such events from spontaneous reports are not meaningful for the purposes of determining increases or decreases in risk with a drug exposure.

Dr. Blume states that analysis of "anecdotal case reports" can "support the existence of a causal link between patients who take Neurontin and suicidality." (Blume Declaration 2008, para 23) Mr. Altman states that PRR analysis can generate a "signal of a safety problem that, when combined with other information, supports the conclusion that Neurontin has the biological capacity to cause patients who take it to commit or attempt suicide." (Altman Declaration 2008, para 27) Both are incorrect. Data mining of spontaneous adverse event reports is used to generate an alert or signal of disproportional reporting (SDR). These terms were selected very carefully, to make it clear that a statistical elevation in reporting rates for a particular drug-event combination is not the same as a "signal", which in pharmacovigilance implies clinical relevance. (Bates A and Edwards IR. 2006) The mere finding of an SDR does not support the conclusion that a drug has the "biological capacity" to cause an event. Nor does it, by itself, support the conclusion that a signal exists. It merely identifies things that are reported at higher rates than expected but without any clinical context or meaning. The finding of an SDR presents, at most, a hypothesis of a "signal" that requires expert clinical case review and interpretation before it can be deemed a signal.

Mr. Altman presents a graph (Altman Declaration 2008. Exhibit C) which compares the proportion of adverse event reports with at least one event within the HLT category "Suicide and self-injurious behavior" as a proportion of something undefined. As he states (Altman Declaration 2008 para 36) the chart shows variation in proportions, with the highest proportion of this HLT category among those who reportedly took the drug for "psychiatric indications." His work clearly illustrates the importance of considering confounding by indication. Suicidal behavior is expected to be higher among patients who are being treated for psychiatric conditions than epilepsy and other indications (e.g. neuroleptic pain), though all have elevated rates of suicide compared to the general population. (Pfizer_Mpatel_0039743, Christensen et al. 2007, Tondo L. et al. 2003)

Mr. Altman compares the number of completed suicide events reported from 1997-2002 (n=8) with the number reported during the first half of 2003 (n=17), which he states a 20-fold increase in the reporting rate of suicide (Altman Declaration 2008 para 21). An increase in reports, in itself, is meaningless. Mr. Altman's methods are unscientific and not appropriate, as he fails to put his numbers within the context of reporting trends during this time period. Thus, his statement without any such context is misleading. While the absolute numbers may be correct, there are a number of flaws in his interpretation.

One flaw in Mr. Altman's comparison is the assumption that the reporting rate for gabapentin was constant during 2003. It actually increased almost 50% from the first to second half of the year. His interpretation of the data is also flawed. For example, reporting rates for any one event are typically calculated as a proportion of all reports for that same drug to correct for changes in reporting volume. Report volume has increased substantially for the AERS database overall and is highly variable for individual drugs. So this "20-fold" or 200% increase in suicide reports must be interpreted within the context of a 360% increase (from 803 to 2896) in adverse event reports that mentioned gabapentin from 1997 to 2003, a 615% increase (from 2.6 to 16 million) in gabapentin prescriptions, and a 174% increase (from 212,978 to 370,898 reports in 1997 and 2003, respectively) in total postmarketing adverse event reports submitted to the FDA during the same period. Thus, a 20-fold increase in the absolute number of suicide reports is less than would be expected within the context of gabapentin adverse event reporting rates. The percentage of reports in which gabapentin was noted as a suspect drug by the reporter remained stable at about one-third (34.4% in 1997 and 34.4% in 2003) of reports which noted the patient had used gapabentin. This suggests that increased number of reports in which the patient used gabapentin tracks with the increased use of the drug over time.

Using published signal detection methods on the FOI-AERS postmarketing database I was able to confirm that Pfizer's conclusions regarding the absence of a signal of suicide for Neurontin were appropriate (Pfizer_Regulatory_001621). Mr. Altman and Dr. Blume did not establish a signal for Neurontin, as they employed methods and interpretations that pharmacoepidemiology experts do not except as reliable or valid.

Mr. Altman declares that there was no notoriety bias in reporting of suicides based solely on the timing of his involvement and the involvement of his law firm. (Altman Declaration 2008, para 30) This ignores the fact that there are many other people and entities reporting events to the FDA. Indeed, the evidence suggests that there was significant "notoriety bias" surrounding the reporting of suicide-related adverse events during the period in question.

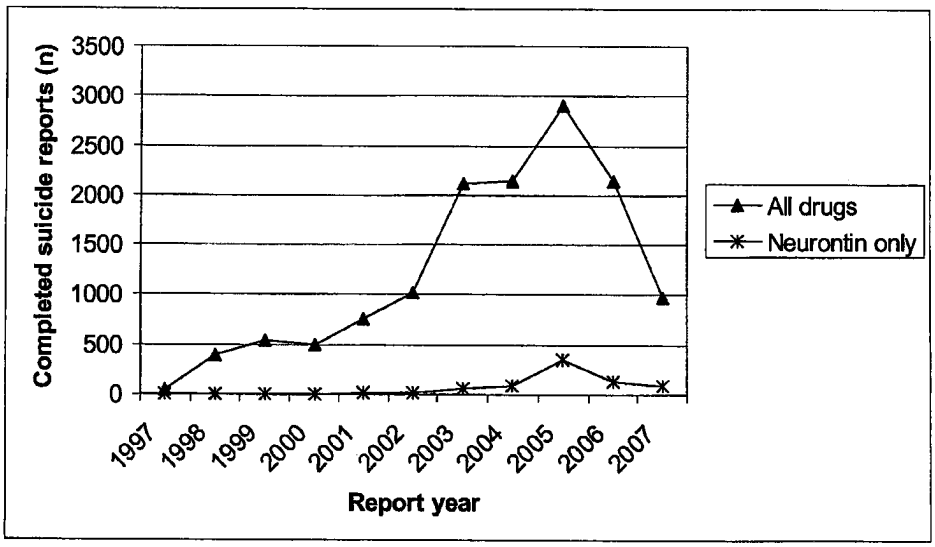


Figure 3. Reports of suicide in AERS for all drugs and for gabapentin

As shown in Figure 3, the number of suicides (completed suicides) reported to the FDA regardless of drug more than doubled from 1027 in 2002 to 2119 reports in 2003, while suicide reports mentioning gabapentin increased 2.3-fold from 40 in 2002 to 92 in 2003. The reporting of completed suicides (all drugs) appears to have peaked in 2005 at 2899 reports. Also during this same time period both the FDA and European regulators issued warnings about a possible link between SSRI antidepressants and suicidal behaviors in children and adolescents in 2003 and FDA held an advisory committee meeting on this subject in early (February 3rd) 2004. Such events are known to stimulate reporting. (FDA Drug Safety Newsletter 2008) Mr. Altman's "analysis" does not account for this increase in the reporting of suicides in the AERS database.

These data are consistent with data recently published by Bridge, et al. (Bridges et al. 2008) In this paper, the authors evaluated data on deaths for which suicide was listed as the underlying cause of death among 10-19 year olds in the National Vital Statistics Systems. They found that although the overall rate of suicide decreased by 5.3% between 2004 and 2005, the rate of suicide during these years was significantly greater than expected based on the 1999-2003 trend. (See Figure 4) Thus, there was a significant increase in suicide rates between 2003 and 2004, at the same time as the increase in overall reports of suicide in AERS (for all drugs).

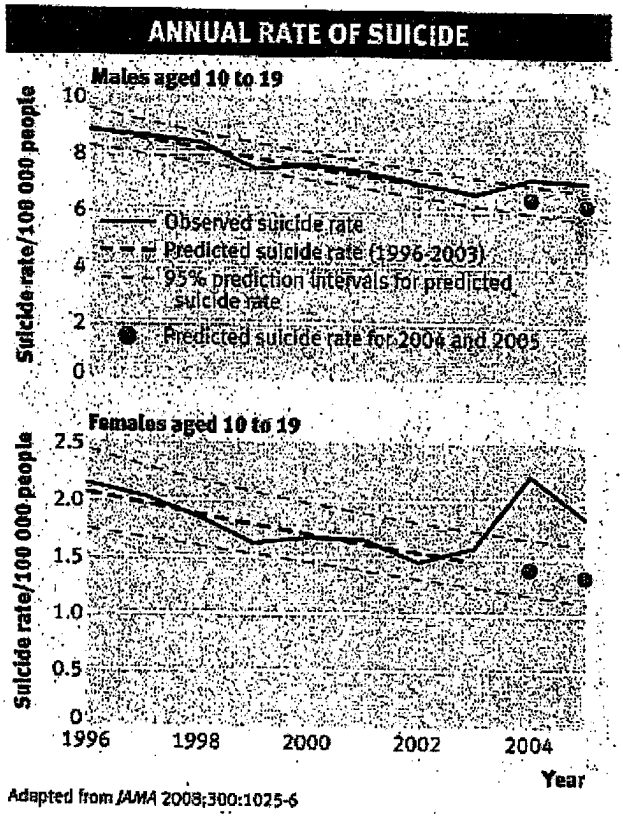


Figure 4. National Trends in suicide from 1996 through 2005. (Bridge JAMA 2005)

The failure of Mr. Altman or Dr. Blume to put any of their "looks" at AER reports for Neurontin within the context of overall reporting trends and national trends in suicide demonstrates the incomplete nature of their analysis and the fallacy of using such data to infer a causal relationship between Neurontin and suicidality.

Mr. Altman further notes (paragraph 29) that the percentage of serious adverse events "exploded" by the 4th quarter of 2002, when "off-label use of Neurontin is at its peaks (sic)." In addition to not defining the term "exploded" in this context he fails to mention that the percentage of serious adverse event reports increased for the entire AERS database, not just Neurontin. Mr. Altman bases his opinion on a flawed interpretation of an increase in serious reports for Neurontin, which is meaningless when looked at properly, within the context of changes to the entire AERS database. This occurred because of a number of processes and policies implemented by FDA beginning in November 1997 with the implementation of the AERS system. This included an abrupt drop in the entering of individual case reports for nonserious events received as periodic reports from manufacturers into AERS and the initiation of a waiver program by which companies could receive waivers from the requirement to submit such individual reports under 21 CFR 314.90.

Mr. Altman does mention the waiver program later in his declaration (Altman 2008 para 38.c.). Based on FDA's Office of Drug Safety's 2002 Annual Report, by September 2002 over 2700 NDA and ANDA's from more than 70 companies had such waivers. (FDA. 2004). Neurontin was one of the drugs which had a waiver. (Confirmed by Counsel upon my inquiry.)

These two factors resulted in a large and rapid drop in individual manufacturer's periodic reports entered into the AERS database, from 73.7% of reports received by FDA in 1997 to 31.1% of reports received in 2003. The proportion of expedited reports, which by their definition are serious, rose from 18.0% of reports in 1997 to 55.4% in 2003. By 2006, the proportion of expedited to periodic reports had almost completed reversed as the FOI-AERS database contained 62.9% expedited reports, 28.7% periodic and 8.3% direct reports. Thus, an increase in the proportion of serious reports for Neurontin mirrors what was happening in the entire AERS database. It can be linked to an intentional process initiated by the FDA to increase the proportion of serious adverse events. Thus, Mr. Altman's assertion that the percentage of serious adverse events "exploded" by the 4th quarter of 2002 is misleading and inaccurate.

Mr. Altman and Dr. Blume conclude that increases in serious adverse event reports in the 4th quarter 2002 were due to off-label uses and is therefore evidence of a link between Neurontin and suicidality. This ignores that the total prescriptions for Neurontin were also increasing during this time period and that reporting of all events and particularly serious adverse events to the FDA were also increasing during this same time period. Also, if there was an increase in the use of this drug for the off-label indication that is associated with an elevated risk of suicidal thoughts and behaviors, such as bipolar disease, then there should be an increase in the reporting of these same events because they are increasing in the background. This is an example of confounding by indication. If it were scientifically valid to jump to the same conclusions that Mr. Altman and Dr. Blume did upon their review of spontaneous reporting trends, then surely the FDA would not have had to spend upwards of three years to conduct a secondary analysis of randomized clinical trial data for 11 antiepileptic drugs and convene an advisory committee meeting to help them interpret the meaning of such analysis. In fact, FDA decided to analyze only randomized placebo-controlled clinical trial ("RCT") data in its analysis of suicidality and AEDs. FDA concluded that post-marketing spontaneous adverse event data were inappropriate for a study of suicidality in the population of patients taking AEDs, as such patients have a high background rate of suicide. As Russell Katz of FDA pointed out at the July 10, 2008 FDA Advisory Committee meeting,

> *"...we had long ago decided that postmarketing data are not the right data to look at, or we don't believe that these sorts of things where there is a high background rate of suicidality so defined in these populations, I think that we have concluded that postmarketing data is uninterpretable, and that is why we went to placebo-controlled trials."* (FDA Transcript 2008, Dr. Katz. Page 103)

Opinion 7. There is no evidence that there are subpopulations at increased risk for suicidality with Neurontin use.

As stated in my initial report, and in this report, Neurontin has been prescribed for conditions that have elevated rates of suicide compared to the general population. In paragraph 36 and Exhibit C of his declaration, Mr. Altman presents data that he purports to show that there is a higher percentage of serious reports for events coded with the HLT "Suicidal and Self-Injurious Behavior" for patients taking Neurontin for psychiatric indications, compared to those taking Neurontin for "anti-epileptic," "other," "unspecified," and "all indications." (Altman K. Declaration 2008, Exhibit C) Once again, Mr. Altman merely presents crude percentage counts and infers that the data show an increased risk for suicide in those patients taking Neurontin for psychiatric indications. Such an inference may not be made from these data. (FDA Drug Safety Newsletter 2008)

Because of the increased risk of suicide among patients for whom Neurontin is generally prescribed, uncontrolled trials and spontaneous reporting data is of no use in making causal inferences or comparisons.

> *"FDA suggests that a comparison of two or more reporting rates be viewed with extreme caution and generally considered exploratory or hypothesis-generating. Reporting rates can by no means be considered incidence rates, for either absolute or comparative purposes."* (FDA Guidance Pharmacovigilance and Pharmacoepidemiology 2008, page 11)

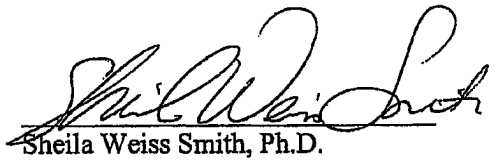
Therefore, the comparisons that Mr. Altman made (Altman, K. Declaration. Para 36 and Exhibit C) provide no evidence of differential risk. They merely show reporting rates of one MedDRA category by the "indication" for which the reporter said that Neurontin was being prescribed. The high proportion of suicidal events is most likely reflecting the elevated risk of suicide and suicidal thoughts among patients with psychiatric conditions. This is called confounding by indication and is a widely recognized problem in the design of pharmacoepidemiology research studies. (Strom BL and Melmon KL. 2005)

> *For any individual case report, it is rarely possible to know with a high level of certainty whether the event was caused by the product. To date, there are no internationally agreed upon standards or criteria for assessing causality in individual cases, especially for events that often occur spontaneously (e.g. stroke, pulmonary embolism). Rigorous pharmacoepidemiologic studies, such as case-control studies and cohort studies with appropriate follow-up, are usually employed to further examine the potential association between a product and an adverse event.* (FDA Guidance Pharmacovigilance and pharmacoepidemiology 2005, page 7)

This helps explain why the FDA did not rely on evaluations of spontaneous reports and asked sponsors to submit only randomized controlled clinical trials to examine the

hypothesis the antiepileptic drugs are associated with suicidal risk. FDA did an analysis by trial indications, the shortcomings of which I discussed earlier. (Opinion 1, section C) Even if you ignore the shortcomings of the FDA analysis, there is no elevated odds ratio for suicidality in any trials except those with an epilepsy indication. (Levenson M. 2008. Figure 8, page 33)

Submitted:



Sheila Weiss Smith, Ph.D.

11/07/08

Date: